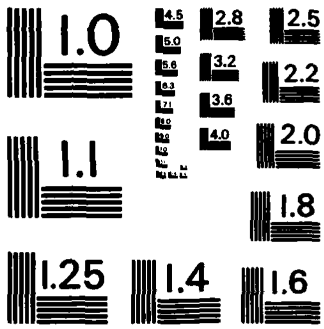
AD-A121 678 AUTOMATIC RECOGNITION AND TRACKING OF OBJECTS(U) TEXAS UNIV AT AUSTIN DEPT OF ELECTRICAL ENGINEERING J K AGGARWAL NOV 81 AFOSR-TR-82-1002 AFOSR-77-3190 1/1

UNCLASSIFIED F/G 17/8 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 82-1002

(4)

AD A121678

Final

Scientific Report for the Period
1 December 1980 - 30 November 1981

Automatic Recognition and Tracking of Objects

Grant AFOSR 77-3190

The University of Texas at Austin
Department of Electrical Engineering
Austin, Texas 78712

Principal Investigator: Professor J. K. Aggarwal

82 11 22 021

SCIENTIFIC REPORT

During the period, December 1, 1980 through November 30, 1981, a major research effort was devoted to the four areas briefly described in the next section of this report. In addition, our group made 5 presentations and prepared 12 papers. Four of the papers were included in conference proceedings, 5 have already appeared in journals and books, and 3 are due to appear in journals. Finally, 2 technical reports were prepared.

1. Motion and Image Differencing:

Differencing operations are used to guide the detection of changes in images due to the motion of objects in a scene. For moving objects of homogeneous grey level, differencing operations applied to pairs of consecutive frames from an image sequence can identify image areas corresponding to a portion of an object in one image but not in the other. If the two positions of an object overlap in the images, the common area does not generate difference picture points. A method for detecting and using these common area points for deriving descriptions of moving polygonal objects was developed earlier and the preliminary results were presented in [A.1]. The analysis program based upon the above method has been expanded to handle more general input scenes and a complete description will appear in [C.1]. Examples discussed in [C.1] include laboratory scenes containing both polygonal and curvilinear objects in a noisy environment. The results of the processing illustrates both the generality and the efficacy of the developed algorithms.

2. Rigid and Jointed Objects:

A new method of interpreting structure from motion has been developed for dynamic scenes containing consistently detectable feature points [A.2,A.3]. Using this technique it is possible to recover the three-dimensional structure (to within a reflection) from a sequence of monocular views of any group of rigidly connected points whose motion satisfies the following motion constraint: FIXED AXIS ASSUMPTION: Every rigid object movement consists of a translation plus a rotation about an axis that is fixed in direction for short periods of time. These fundamental results are discussed in [B.1], while a complete description with formal proofs is being prepared and will appear in [C.2]. These results have been applied to several sets of data including a person swinging a baseball bat and a person walking. The results on the connection structure and rigid part lengths are mixed.

3. Multiple views and occluding contours:

The problem of determining the three dimensional description of an object from multiple views of the occluding contour is under investigation as reported in [A.4]. The fundamental objective of this work is to lessen the dependence on single feature point detection and token correspondence, and to develop more descriptive representations of three-dimensional objects. Instead of attempting to isolate individual points of interest in the images, our aim is to apply simple processes to detect the occluding contour, i.e., the silhouette, of the object. The system uses several silhouettes of an object, i.e., the occluding contours from several viewpoints, to form bounding volumes for the object. New information about the object surface is accumulated from the image sequence and spatial models are created so that explicit specifications of the imaging coordinate system are required. Under orthographic projection, the direction of the view line of the imaging system is of dominant importance, while the optical origin may, in some cases, remain unspecified.

Each point on a contour determines a projection line that is tangent to the object surface at some point, or set of points. The possible positions of those tangent points along the projection line are constrained by the boundaries of each of the remaining contours, and as new views are acquired, further refinements of that estimate are specified. The result is a volume within which the object must lie. If any two of the view lines are not parallel then the resulting volume is closed and bounded. It is important to observe here that the contours need not be searched for corresponding points which identify the same feature on the actual object surface. The necessary correspondence information is provided by the orientations of the view lines. This work is appropriate to industrial automation applications. For example, selecting one of several parts on a conveyor using the views taken from several fixed cameras.

4. Three-Dimensional Description of Objects:

The scheme of Section 3 requires a representation which captures the necessary surface details and remains flexible enough to facilitate the continual refinement that is fundamental to the process. A complete review of the available techniques for characterizing three-dimensional objects has been undertaken and the results will appear in [C.3]. None of the methods reviewed were suitable for the system of Section 3, so a "volume segment" representative was defined and implemented [D.1]. The representation comprises a set of parallel segments that define the bounding volume in a globally defined x-y-z coordinate system.

The primary advantage of this representation is that the process of determining whether an arbitrary point is within the surface boundary consists of a simple search of three ordered lists: select a "plane" by the z-coordinate; select a "line" by the x-coordinate; and finally, check for inclusion of the y-coordinate in a segment. This structure can also provide a fairly succinct representation, particularly for objects that are elongated in the direction of the y-axis.

5. Intensity and Range Information:

A laser ranging instrument can determine the distance from the instrument to a given point in a scene along a straight line. A complete range image may be constructed by scanning the visual field in a raster fashion. Each pixel value in the range image along with row and column of the image gives the spherical coordinates of its corresponding scene point. The (x,y,z) Cartesian coordinates can be computed by simple coordinate transformation. Thus the surfaces in the scene are described by their Cartesian 3-D coordinates. At the same time, the ranging instrument gives the intensity at each pixel. Thus the ranging instrument gives essentially two distinct pieces of information - range and intensity.

We have considered a variety of issues based upon the problems of combining range and intensity information, and of determining the surfaces from range and intensity information. Methods for extracting planar surfaces from the range information were also considered. These methods will not only give the horizontal and vertical surfaces but will develop the orientation and position of slated surfaces. Intensity data will also be examined to determine "edges." The next step and probably the most difficult step involves the combining of the two knowledge sources to obtain a single segmentation of the scene. This investigation is continuing and a report documenting preliminary results is under preparation. In addition, we shall address these issues by applying object model restrictions and assumptions concerning the lighting conditions.

6. Related Research:

Sections 1 through 4 describe projects concerned with various aspects of dynamic scene analysis. These projects are part of a continuing research effort to gain an understanding of the fundamental principles of time-varying imagery and to develop appropriate methods to process such imagery. This broad interest in dynamic scene analysis is reflected in the completed papers [B.2, B.3, B.4].

Another paper [B.5] that appeared in this period described the results obtained from the analysis of the chromatic images yielded by our flying-spot-scanner. In this case the segmentation of single color transparencies was the primary concern.

PRESENTATIONS AND PUBLICATIONS

A. Presentations

1. S. Yalamanchili and J. K. Aggarwal, "Motion and Image Differencing," Proceedings of the IEEE-Pattern Recognition & Image Processing Conference, Dallas, TX, August 1981, pp.211-216.

2. J. A. Webb and J. K. Aggarwal, "Visual Interpretation of the Motion of Objects in Space, "Proceedings of the IEEE-Pattern Recognition & Image Processing Conference, Dallas, TX, August 1981, pp. 516-521.

3. J. A. Webb and J. K. Aggarwal, "Structure from Motion of Rigid and Jointed Objects," Proceedings of the Seventh International Joint Conference on Artificial Intelligence, Vancouver, Canada, August 1981, pp. 686-691.

4. W. N. Martin and J. K. Aggarwal, "Occluding Contours in Dynamic Scenes," Proceedings of the IEEE-Pattern Recognition & Image Processing Conference, Dallas, TX, August 1981, pp. 189-192.

5. J. K. Aggarwal, "Data, Image and Signal Processing vs. Artificial Intelligence," ASSP Workshop on Two-Dimensional Signal Processing, Oct. 5-7, 1981, New Paltz, N. Y. No Proceedings were published.

6. J. K. Aggarwal, "Segmentation and Range Information," ASSP

Workshop on Two-Dimensional Signal Processing, Oct. 5-7, 1981, New Paltz, N. Y. No Proceedings were published.

B. Papers

1. J. A. Webb and J. K. Aggarwal, "Visually Interpreting the Motion of Objects in Space," IEEE Computer Society Computer, August 1981, pp. 40-46.

2. L. S. Davis, W. N. Martin and J. K. Aggarwal, "Correspondence Processes in Dynamic Scene Analysis," IEEE Proceedings, Vol. 69, No. 5, May 1981, pp. 562-572.

3. W. N. Martin and J. K. Aggarwal, "Analyzing Dynamic Scenes Containing Multiple Moving Objects," In Image Sequence Analysis, T. S. Huang, ed., Springer-Verlag, 1981, pp. 355-380.

4. W. N. Martin and J. K. Aggarwal, "Occlusion in Dynamic Scene Analysis," in Digital Image Processing and Analysis, J. C. Simon and R. Haralick, eds., D. Reidel Publishing Co., 1981.

5. A. Sarabi and J. K. Aggarwal, "Segmentation of Chromatic Images," Pattern Recognition, Vol. 13, No. 6, pp. 417-427, 1981.

C. Papers Prepared - To Appear

1. S. Yalamanchili, W. N. Martin and J. K. Aggarwal, "Extraction of Moving Object Descriptions via Differencing," to

appear in Computer Graphics and Image Processing.

2. J. A. Webb and J. K. Aggarwal, "Visual Interpretation of the Motion of Rigid and Jointed Objects," to appear in Artificial Intelligence.

3. J. K. Aggarwal, L. S. Davis, W. N. Martin and J. W. Roach, "Survey: Representation Methods for Three-Dimensional Objects," to appear in Progress in Pattern Recognition, L. N. Kanal and A. Rosenfeld, eds., North-Holland.

D. Reports

1. J. A. Webb and J. K. Aggarwal, "Visually Interpreting the Motion of Objects in Space," Laboratory for Image and Signal Analysis, TR-81-3.

2. W. N. Martin and J. K. Aggarwal, "Analyzing Dynamic Scenes," Laboratory for Image and Signal Analysis, TR-81-5.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR- 82-1002	2. GOVT ACCESSION NO.: AD-A121618
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): AUTOMATIC RECOGNITION AND TRACKING OF OBJECTS	5. TYPE OF REPORT & PERIOD COVERED: FINAL 01 DEC 76 to 30 NOV 81
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Professor J.K. Aggarwal	8. CONTRACT OR GRANT NUMBER(s): AFOSR- 77-3190
9. PERFORMING ORGANIZATION NAME AND ADDRESS: The University of Texas at Austin, Austin, Texas 78712	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 2305/B1 61102F
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research, Building #410, Bolling AFB, Washington, D.C. 20332	12. REPORT DATE: Nov 1981
	13. NUMBER OF PAGES: 09
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The broad program of research consisted of multisensor image understanding including integration of information from multiple sources, the tracking of objects in a sequence of images in real time, together with the estimation of motion parameters, the characterization of the descriptions and invariances of objects, and the registration of objects. Intensity, color, and range were used in the segmentation of scenes, and the extraction and identification of general areas of interest in the scenes. Shape descriptors provide structural descriptions of objects and lead to recognition of

OVER

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) objects and background components. Motion analysis is instrumental in tracking and prediction of the movement of objects. The analysis and understanding of the structure and motion of three-dimensional space from a sequence of two-dimensional images in real time is the fundamental goal of the present investigation.